

*

Auteur: Harald Merckelbach
Bron: Skepter 15(4), december 2002



*

Rekenen met Gigerenzer

boekbespreking

Statistisch onbegrip is niet voorbehouden aan leken. Ook medici en juristen hebben er last van, soms met fatale gevolgen. Kan daar iets aan gedaan worden?

Stel dat er een nieuwe behandeling wordt ontdekt voor een bepaalde aandoening. Maar stel ook dat de kans 7% is dat patiënten binnen 5 jaar aan deze nieuwe behandeling overlijden. Zou u uw patiënten deze behandeling aanraden? Toen de Amerikaanse psycholoog Amos Tversky die vraag aan artsen voorlegde, merkte hij dat zij zeer terughoudend waren met hun aanbevelingen. Die terughoudendheid maakte plaats voor groot enthousiasme indien de artsen het probleem als volgt kregen voorgeschoteld: 93% van de patiënten die de behandeling ondergaat, zal de 5 jaren overleven. Met een lange reeks van dit soort gedachte-experimenten lieten Tversky en zijn collega Daniel Kahneman zien dat artsen en allerlei andere specialisten, maar trouwens ook gewone mensen slecht zijn in het interpreteren van kansen. Daarmee stonden Tversky en Kahneman aan de wieg van wat thans een omvangrijke tak van sport in de psychologie is: de psychologie van de cognitieve vertekeningen ('cognitive biases'). Deze vorm van psychologie rekt het tot haar hoofdtaak om te laten zien hoe irrationeel mensen zich gaan gedragen als ze keuzes moeten maken in situaties die een taxatie van kansen vereisen.

Het onderzoek naar 'cognitive biases' begint inmiddels op grote schaal haar weg te vinden naar disciplines waar kansen een sleutelrol spelen. Te denken valt aan economische vraagstukken (de kans dat de aex graadmeter in elkaar zakt), juridische kwesties (de kans dat de verdachte schuldig is) en natuurlijk medische problemen (de kans dat een behandeling de patiënt van een wisse dood redt).[1] Vanwege de uitstraling die het heeft naar andere vakgebieden, mag er volgens de Italiaanse psycholoog Palmarini (1996) zo onderhand wel een Nobelprijs worden uitgereikt voor dit pionierswerk. (Inmiddels heeft Kahneman de Nobelprijs voor economie in ontvangst genomen (zie kader) – red.)

Tversky en Kahneman en de onderzoekers die in hun voetsporen traden, exporteerden niet alleen hun

onderzoekresultaten, maar ook een weinig opbeurend mensbeeld naar belerende disciplines. Mensen, zo luidt de boodschap van Tversky & Kahneman, zijn nauwelijks in staat om op een rationele manier keuzes te maken. De Duitse psycholoog Gigerenzer verzet zich met hand en tand tegen die boodschap. Hij meent dat het veelal een kwestie van presentatie is, waarover straks meer.



Gigerenzer erkent dat mensen grote moeite hebben met het interpreteren van kansinformatie. Hij haalt in zijn boek het voorbeeld aan van een Duitse krant die aan haar lezers vroeg wat '40%' betekent: een kwart, 4 uit 10 of elke 40ste persoon? Zo'n 33% van de lezers (bijna 1 op de 3 dus) koos een fout alternatief. Dat lijkt iets te zeggen over 's mensen ongecijferdheid, maar, aldus Gigerenzer, we moeten bedenken dat de geschiedenis van kansen en statistieken een korte is. De wiskunde is begonnen er in de tweede helft van de 17de eeuw over na te denken, de door Napoleon ingestelde *bureaux de statistique* voorzagen het publiek pas vanaf het midden van de 19de eeuw mondjesmaat van statistieken en dat de weerman probabilistische termen hanteert ('de kans op regen is 50%') is iets van de laatste 10 jaar.

Dat kansinformatie in weerberichten opduikt en dan een misleidende uitwerking kan

Samen met zijn collega Amos Tversky (die in 1996 overleed) ontwikkelde de psycholoog Daniel Kahneman de zogenoemde *prospect theory*, waarover hij in 1979 een artikel publiceerde in het gerenommeerde tijdschrift *Econometrica*. Als het destijds in een psychologisch tijdschrift was verschenen, zou hij de Nobelprijs voor economie waarschijnlijk niet hebben ontvangen. De *prospect theory* gaf de aanzet tot een nieuwe benadering binnen de economische wetenschap, waarbij men door middel van experimenten onderzoekt hoe financiële beslissingen afwijken van het rationele keuzemodel van de traditionele economische theorie.

Economen gingen uit van de veronderstelling dat mensen op basis van de beschikbare informatie vaststellen welk alternatief de hoogste verwachte waarde (utiliteit) heeft, maar in werkelijkheid kiest men lang niet altijd het meest gunstige alternatief. Wanneer mensen winst kunnen boeken, nemen ze bij voorkeur het zekere voor het onzekere: ze incasseren liever 3000 euro dan dat ze 80 procent kans hebben om 4000 euro op te strijken. Dit verandert echter wanneer het niet om winst maar om verlies gaat. Mensen nemen meer risico om verlies te vermijden dan om winst te verkrijgen, waardoor ze soms meer verliezen dan nodig is. Een verlies weegt voor hun gevoel veel zwaarder dan een even grote winst. Het hangt daarbij van het subjectief gekozen referentiepunt af of een bepaalde uitkomst als een winst of als een verlies wordt geïnterpreteerd.

Een ander interessant verschijnsel, het *endowment effect*, kan onder meer worden waargenomen wanneer je in een groep studenten aan de helft van de aanwezigen koffiemokken uitdeelt en aan de andere helft plakken chocolade die ongeveer even duur zijn. Vraag vervolgens wie zijn mok wil ruilen tegen een

hebben is tot daaraan toe.[2] Zorgwekkender is de opmars van kansinformatie in het medische en juridische domein. Daar komt die informatie in handen te liggen van experts die tijdens hun opleiding zich de nodige moeite zullen hebben getroost om twee dingen te vermijden: statistiek en

stuk chocolade, of omgekeerd. Wanneer de objecten willekeurig zijn uitgedeeld, dan lijkt het aannemelijk dat ongeveer de helft van de studenten liever iets anders heeft. In werkelijkheid blijkt vaak dat niet meer dan tien procent wil ruilen. De overgrote meerderheid prefereert de status quo. Uit dit soort experimenten blijkt dat mensen gewoonlijk veel meer geld willen hebben om een voorwerp af te staan dan ze zouden willen betalen om het te verwerven. (red.)

psychologie. Statistiek heb je echter hard nodig om te begrijpen wat een kans is en psychologisch inzicht is nooit weg als je dat begrip vervolgens wilt overdragen aan patiënten of rechters. In zijn boek staat Gigerenzer uitgebreid stil bij medische en juridische voorbeelden waarbij statistisch onbegrip fatale gevolgen had. De strekking van die voorbeelden is steeds dat dokters en juristen een diepe afkeer hebben van onzekerheid en zodoende ertoe neigen om absolutistische interpretaties te omarmen. Zo kan het gebeuren dat in Duitsland per jaar zo'n 100.000 vrouwen onnodig invasieve chirurgie ondergaan nadat hun mammografie een ongunstige uitslag heeft opgeleverd. In dit geval nemen medici nogal eens aan dat de kans op borstkanker gegeven een ongunstige testuitslag (p borstkanker/test) gelijk is aan de kans op een ongunstige testuitslag gegeven borstkanker (p test/borstkanker). Omdat die laatste kans ver in de 90% loopt, terwijl die eerste kans rond de 7% cirkelt en bovendien borstkanker bij veel minder dan de helft van de vrouwen voorkomt, onderschatten specialisten de kans op fout-positieve diagnoses (borstkanker diagnosticeren bij gezonde vrouwen) als men de test bij wijze van screening gaat toepassen op grote populaties.

Te hoog cholesterol

Het negeren van fout-positieve uitslagen is ook aan de orde in het voorbeeld van de 7 bloeddonoren uit Florida die de hand aan zichzelf sloegen. Dat gebeurde nadat een elisa-test zou hebben uitgewezen dat zij met hiv waren geïnfecteerd. De specialisten vergaten erbij te vertellen dat zo'n test een ruime marge (enkele tientallen procenten) van fout-positieve uitslagen kent en dat het dus heel goed denkbaar is dat een gezond iemand toch een alarmerende elisa-uitslag heeft. Het verhaal van deze bloeddonoren speelt zich af in de jaren 1980 en je zou dan ook verwachten dat op dit moment specialisten met een goed verhaal klaar staan als mensen zich willen laten onderzoeken op hiv. Gigerenzer en een van zijn studenten namen in Duitsland de proef op de som. De student meldde zich bij 20 verschillende gezondheidscentra voor een hiv-test. Hij vroeg telkens naar de kans op een fout-positieve uitslag, maar werd door het merendeel van de specialisten afgescheept met de mededeling dat zo'n uitslag nooit voorkomt. Artsen staan niet graag stil bij fout-positieve uitslagen. Die zeggen namelijk iets over de onbetrouwbaarheid van hun tests en ondergraven daarmee hun professionele status.[3]

Voor belanghebbenden kan statistische naïviteit een godsgeschenk zijn. Zo

maakten Engelse onderzoekers een aantal jaren geleden met veel bombarie bekend dat het cholesterolverlagend middel pravastatin de kans op een fatale hartaanval met 22% reduceert. Wie zou deze onderzoekers niet willen subsidiëren? En hoe dom moet je zijn om pravastatin niet te slikken? Gigerenzer rekt in zijn boek voor hoe de onderzoekers aan het percentage van 22% kwamen. Ze hadden twee groepen van telkens 1000 patiënten met een te hoog cholesterolgehalte. De ene groep kreeg pravastatin, de andere groep slikte een placebo en beide groepen werden 5 jaar lang gevolgd. In de groep met pravastatin overleden in die periode 32 mensen terwijl in de placebogroep de teller pas op 41 bleef steken. In relatieve termen is de winst die met het middel te boeken valt 9/41 en derhalve 22%. In absolute termen uitgedrukt, is die winst echter 9 per 1000, dus nog geen 1%. Overigens is het onderscheid tussen 32 en 41 statistisch niet significant.

Vanwege hun aversie tegen onzekerheid hebben artsen de neiging om fout-positieven te onderschatten en onderzoeksresultaten te overschatten. Diezelfde neiging treft men aan bij rechters en officieren van justitie. Op geleide van wat de experts – vaak dokters – hun vertellen, koesteren zij een blind vertrouwen in de bewijswaarde van vingerafdrukken, bloedtypen en DNA. Vooral DNA-bewijs heeft inmiddels de status van gouden standaard verworven. Het ontzag voor dit bewijsmiddel is zo groot dat bijvoorbeeld een kamerlid als Boris Dittrich pleit voor het afschaffen van de verjaringstermijnen bij moord en doodslag. Het idee hierachter is dat ook DNA-bewijs dat pas na jaren opduikt rechtstreeks en onomstotelijk leidt naar dader. Maar hoe waterdicht is dit bewijsmiddel? Hoeveel fout-positieve uitkomsten genereren de door justitie ingeschakelde DNA-laboratoria? Gigerenzer benadrukt dat we akelig weinig weten over de kans op fout-positieven bij DNA-evidentie, maar dat er toch redenen zijn om aan te nemen dat die soms in orde van grootte van 1 op de 200 is.[4] Dat betekent dat er één onschuldige verdachte dikke pech heeft op elke 200 gevallen waarin er een 'DNA match' heet te zijn tussen biologisch materiaal (sperma, een haar) afkomstig van de plaats delict en de verdachte.

Prosecutor's fallacy

Gigerenzer wijst erop dat bij DNA bewijs vaak nog een ander probleem gaat spelen, namelijk dat van de 'prosecutor's fallacy'. In de rechtszaal geven DNA-specialisten een cijfermatige waardering aan een overeenkomst tussen het DNA van de verdachte en dat van op de plaats delict veilig gestelde biomaterialen. Die waardering is de 'random match probability': de kans dat het DNA van een willekeurig individu uit de populatie per toeval overeenkomt met het biomateriaal van de plaats delict. Stel – zoals in een Schots geval aan de orde was – dat de specialist de 'random match probability' op 1 uit de 49.000 mensen schat. Officieren van justitie leggen die kans dan als volgt uit: het DNA van de verdachte vertoont een gelijkenis met dat van het sperma of het bloed dat op de plaats van de moord werd gevonden; de kans dat dit toeval en de verdachte derhalve onschuldig is, bedraagt 1/49000 en de kans dat de verdachte schuldig is wordt daarmee 48999/49000. Dat is een getal waarmee je verdachten behoorlijk

snel achter de tralies krijgt. Bij de 'prosecutor's fallacy' wordt echter de kans op een 'random match' (p match) verward met de kans dat déze verdachte onschuldig is en toch een match laat zien (p onschuldig/match). Wat die verwarring betekent, wordt duidelijk als men zich probeert voor te stellen dat het DNA in een stad met 100.000 volwassenen is gescreend. Met de 'random match probability' in het Schotse geval levert dat ongeveer 2 treffers op en dan wordt de kans op onschuld voor elk van de twee verdachten 50%. 'Random match probabilities' kunnen trouwens behoorlijk groot worden voor verwanten van de verdachte. In het eerder aangehaalde Schotse geval bleek de kans 1 op 16 te zijn dat DNA van een familielid van de verdachte een treffer met het biomateriaal vertoonde. Het toeval wilde dat de verdachte 5 broers had.

Death & Taxes

Valt er iets te doen aan de misleidende uitwerking die kansinformatie op dokters, juristen en leken heeft? Onderzoek in de traditie van Tversky en Kahneman suggereert vaak dat we hier te maken hebben met inherente beperkingen van het menselijk denkvermogen. Gigerenzer is aanmerkelijk optimistischer en stelt zich op het standpunt dat het gaat om reparabele afwijkingen. De reparatie bestaat hieruit dat je nooit kansinformatie accepteert en eist dat dit type informatie wordt geherformuleerd in termen van natuurlijke frequenties. De ogenschijnlijk geringe kans van 7% op fout-positieve uitkomsten bij pakweg mammografie blijkt dan in een screening met 1000 gezonde vrouwen 70 onterechte treffers op te leveren. Dat type informatie begrijpen mensen aanmerkelijk beter, ondanks het feit dat het de aangename illusie van een waterdichte test doorgaans logenstraft.

Sceptici maken zich graag vrolijk over het obscurantisme van alternatieve autodidacten. Zulke autodidacten zijn echter aanmerkelijk makkelijker te herkennen en daarom onschuldiger dan de misleidende informatie waar Gigerenzer het over heeft. Zijn boek laat zich dan ook het beste lezen met pen en papier in de aanslag. Er moet gerekend worden, maar wie bereid is om dat te doen ziet dat Gigerenzer gelijk heeft als hij Franklins motto tot het zijne maakt: 'In this world, nothing is certain but death and taxes'.

Noten

[1] Een aardig overzicht van hun invloed op andere disciplines biedt de door Kahneman en Tversky (2000) geredigeerde en meer dan 800 pagina's tellende bundel.

[2] Gigerenzer geeft in zijn boek het vermakelijke voorbeeld van de weerman die voor zowel zaterdag als zondag een 50% kans op regen voorspelde en afsloot met de constatering dat het dus zeker ging regenen in het weekend.

[3] De veronachtzaming van fout-positieve uitslagen kan in de rechtszaal een sterke coalitie tussen de medisch expert en de officier van justitie opleveren. De

medisch expert neemt dan stellige uitspraken voor zijn rekening over het slachtoffer de symptomen wijzen op een trauma – en dat wordt door de officier als bewijs tegen de verdachte ingebracht. Zie Merckelbach en Rassin (2001).

[4]. Zie voor een meer uitgebreide discussie van dit punt: Thompson (1997). Als fout-positieve uitkomsten een echt probleem zijn bij 'DNA fingerprinting', dan zou je verwachten dat de eerste gevallen van op basis van DNA bewijs veroordeelde onschuldigen inmiddels zijn opgedoken. Zulke gevallen bestaan inderdaad en Thompson geeft er een hilarische beschrijving van.

Literatuur

Gigerenzer, Gerd (2002). *Calculated risks: How to know when numbers deceive you*. New York: Simon & Schuster (310 blz.), 34,50 euro.

Kahneman, D. & Tversky, A. (2000). *Choice, values, and frames*. New York: University Press.

Merckelbach, H. & Rassin, E. (2000). Interdisciplinarity in medical and psychological forensic expertise: Problems and solutions. In J.F. Nijboer & W.J.J.M. Sprangers (Eds.). *Harmonisation in forensic expertise: An inquiry into the desirability of and opportunities for international standards*. Amsterdam, Thela-thesis, pp. 145-157.

Palmarini, Massimo Piatelli (1996). *Onvermijdelijke illusies*. Utrecht: Spectrum.

Thompson, W.C. (1997). The national research council's second report on forensic DNA evidence: A critique. *Jurimetrics*, 37, 405-422.

[Harald Merckelbach is als hoogleraar psychologie verbonden aan de Universiteit Maastricht.](#)

[HOMEPAGE SKEPSIS](#)

